0590
0320

#11

OIPE

ENTERED

P.5

RAW SEQUENCE LISTING   DATE: 03/25/2002
PATENT APPLICATION: US/09/777,430A   TIME: 10:43:00

Input Set : A:\04912-1.txt
Output Set: N:\CRF3\03252002\I777430A.raw

```
 3 <110> APPLICANT: Lyamichev, Victor
 4         Skrzypczynski, Zbigniew
 5         Allawi, Hatim T.
 6         Wayland, Sarah R.
 7         Takova, Tsetska
 8         Neir, Bruce P.
10 <120> TITLE OF INVENTION: Charge Tags and the Separation of Nucleic Acid Molecules
12 <130> FILE REFERENCE: FORS-04912
14 <140> CURRENT APPLICATION NUMBER: 09/777,430A
15 <141> CURRENT FILING DATE: 2001-02-06
17 <160> NUMBER OF SEQ ID NOS: 85
19 <170> SOFTWARE: PatentIn version 3.1
21 <210> SEQ ID NO: 1
22 <211> LENGTH: 21
23 <212> TYPE: DNA
24 <213> ORGANISM: Artificial Sequence
26 <220> FEATURE:
27 <223> OTHER INFORMATION: Synthetic
29 <220> FEATURE:
30 <221> NAME/KEY: misc_feature
31 <222> LOCATION: (1)..(2)
32 <223> OTHER INFORMATION: The residues at these positions are amino-modified bases,
bearing
33         a positively charged Cy3 dye.
36 <400> SEQUENCE: 1
37 tcttttcacc agcgagacgg g                                        21
40 <210> SEQ ID NO: 2
41 <211> LENGTH: 23
42 <212> TYPE: DNA
43 <213> ORGANISM: Artificial Sequence
45 <220> FEATURE:
46 <223> OTHER INFORMATION: Synthetic
48 <400> SEQUENCE: 2
49 tattgggcgc cagggtggtt ttt                                      23
52 <210> SEQ ID NO: 3
53 <211> LENGTH: 30
54 <212> TYPE: DNA
55 <213> ORGANISM: Artificial Sequence
57 <220> FEATURE:
58 <223> OTHER INFORMATION: Synthetic
60 <400> SEQUENCE: 3
61 cagggtgaag ggaagaagaa agcgaaaggt                               30
64 <210> SEQ ID NO: 4
65 <211> LENGTH: 30
```

RAW SEQUENCE LISTING   DATE: 03/25/2002
PATENT APPLICATION: **US/09/777,430A**   TIME: 10:43:00

Input Set : **A:\04912-1.txt**
Output Set: **N:\CRF3\03252002\I777430A.raw**

```
 66 <212> TYPE: DNA
 67 <213> ORGANISM: Artificial Sequence
 69 <220> FEATURE:
 70 <223> OTHER INFORMATION: Synthetic
 72 <400> SEQUENCE: 4
 73 caggggggaag ggaagaagaa agcgaaaggt                          30
 76 <210> SEQ ID NO: 5
 77 <211> LENGTH: 28
 78 <212> TYPE: DNA
 79 <213> ORGANISM: Artificial Sequence
 81 <220> FEATURE:
 82 <223> OTHER INFORMATION: Synthetic
 84 <400> SEQUENCE: 5
 85 cacgaattcc gaggcgatgc ttccgctc                             28
 88 <210> SEQ ID NO: 6
 89 <211> LENGTH: 30
 90 <212> TYPE: DNA
 91 <213> ORGANISM: Artificial Sequence
 93 <220> FEATURE:
 94 <223> OTHER INFORMATION: Synthetic
 96 <400> SEQUENCE: 6
 97 tcgacgtcga ctaacccttg gcggaaagcc                           30
100 <210> SEQ ID NO: 7
101 <211> LENGTH: 23
102 <212> TYPE: DNA
103 <213> ORGANISM: Artificial Sequence
105 <220> FEATURE:
106 <223> OTHER INFORMATION: Synthetic
108 <400> SEQUENCE: 7
109 gcatcgcctc ggaattcatg gtc                                  23
112 <210> SEQ ID NO: 8
113 <211> LENGTH: 836
114 <212> TYPE: PRT
115 <213> ORGANISM: Thermus thermophilus
117 <400> SEQUENCE: 8
119 Met Asn Ser Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val
120 1               5                   10                  15
123 Leu Leu Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu
124             20                  25                  30
127 Lys Gly Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly
128         35                  40                  45
131 Phe Ala Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala
132     50                  55                  60
135 Val Phe Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala
136 65                  70                  75                  80
139 Tyr Glu Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro
140             85                  90                  95
143 Arg Gln Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr
144             100                 105                 110
```

RAW SEQUENCE LISTING   DATE: 03/25/2002
PATENT APPLICATION: US/09/777,430A   TIME: 10:43:00

Input Set : A:\04912-1.txt
Output Set: N:\CRF3\03252002\I777430A.raw

```
147 Arg Leu Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu
148         115             120             125
151 Ala Lys Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala
152         130             135             140
155 Asp Arg Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His
156 145             150             155             160
159 Pro Glu Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly
160             165             170             175
163 Leu Arg Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro
164         180             185             190
167 Ser Asp Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu
168         195             200             205
171 Lys Leu Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu
172         210             215             220
175 Asp Arg Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu
176 225             230             235             240
179 Glu Asp Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu
180             245             250             255
183 Pro Leu Glu Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly
184             260             265             270
187 Leu Arg Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu
188         275             280             285
191 Phe Gly Leu Leu Glu Ala Pro Ala Pro Leu Glu Glu Ala Pro Trp Pro
192         290             295             300
195 Pro Pro Glu Gly Ala Phe Val Gly Phe Val Leu Ser Arg Pro Glu Pro
196 305             310             315             320
199 Met Trp Ala Glu Leu Lys Ala Leu Ala Ala Cys Arg Asp Gly Arg Val
200             325             330             335
203 His Arg Ala Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val
204         340             345             350
207 Arg Gly Leu Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly
208         355             360             365
211 Leu Asp Leu Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu
212         370             375             380
215 Asp Pro Ser Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly
216 385             390             395             400
219 Glu Trp Thr Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu
220             405             410             415
223 His Arg Asn Leu Leu Lys Arg Leu Glu Gly Glu Glu Lys Leu Leu Trp
224         420             425             430
227 Leu Tyr His Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met
228         435             440             445
231 Glu Ala Thr Gly Val Arg Arg Asp Val Ala Tyr Leu Gln Ala Leu Ser
232         450             455             460
235 Leu Glu Leu Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg
236 465             470             475             480
239 Leu Ala Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg
240             485             490             495
243 Val Leu Phe Asp Glu Leu Arg Leu Pro Ala Leu Gly Lys Thr Gln Lys
```

Input Set : A:\04912-1.txt
Output Set: N:\CRF3\03252002\I777430A.raw

```
244              500                505                510
247 Thr Gly Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu
248          515                520                525
251 Ala His Pro Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys
252      530                535                540
255 Leu Lys Asn Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Arg
256 545                550                555                560
259 Thr Gly Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly
260              565                570                575
263 Arg Leu Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr
264          580                585                590
267 Pro Leu Gly Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp
268          595                600                605
271 Ala Leu Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala
272      610                615                620
275 His Leu Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys
276 625              630                635                640
279 Asp Ile His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu
280              645                650                655
283 Ala Val Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly
284          660                665                670
287 Val Leu Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile
288          675                680                685
291 Pro Tyr Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe
292      690                695                700
295 Pro Lys Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys
296 705              710                715                720
299 Arg Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp
300              725                730                735
303 Leu Asn Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala
304          740                745                750
307 Phe Asn Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala
308          755                760                765
311 Met Val Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu
312      770                775                780
315 Leu Gln Val His Asp Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala
316 785                790                795                800
319 Glu Glu Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro
320          805                810                815
323 Leu Ala Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu
324          820                825                830
327 Ser Ala Lys Gly
328          835
331 <210> SEQ ID NO: 9
332 <211> LENGTH: 2511
333 <212> TYPE: DNA
334 <213> ORGANISM: Thermus thermophilus
336 <400> SEQUENCE: 9
337 atgaattccg aggcgatgct tccgctcttt gaacccaaag gccgggtcct cctggtggac        60
```

```
339 ggccaccacc tggcctaccg caccttcttc gccctgaagg gcctcaccac gagccggggc    120
341 gaaccggtgc aggcggtcta cggcttcgcc aagagcctcc tcaaggccct gaaggaggac    180
343 gggtacaagg ccgtcttcgt ggtctttgac gccaaggccc cctccttccg ccacgaggcc    240
345 tacgaggcct acaaggcggg gagggccccg acccccgagg acttcccccg gcagctcgcc    300
347 ctcatcaagg agctggtgga cctcctgggg tttacccgcc tcgaggtccc cggctacgag    360
349 gcggacgacg ttctcgccac cctggccaag aaggcggaaa aggaggggta cgaggtgcgc    420
351 atcctcaccg ccgaccgcga cctctaccaa ctcgtctccg accgcgtcgc cgtcctccac    480
353 cccgagggcc acctcatcac cccggagtgg ctttgggaga agtacggcct caggccggag    540
355 cagtgggtgg acttccgcgc cctcgtgggg gacccctccg acaacctccc cggggtcaag    600
357 ggcatcgggg agaagaccgc cctcaagctc ctcaaggagt ggggaagcct ggaaaacctc    660
359 ctcaagaacc tggaccgggt aaagccagaa aacgtccggg agaagatcaa ggcccacctg    720
361 gaagacctca ggctctcctt ggagctctcc cgggtgcgca ccgacctccc cctggaggtg    780
363 gacctcgccc aggggcggga gcccgaccgg gaggggctta gggccttcct ggagaggctg    840
365 gagttcggca gcctcctcca cgagttcggc ctcctggagg cccccgcccc cctggaggag    900
367 gcccccctggc ccccgccgga aggggccttc gtgggcttcg tcctctcccg ccccgagccc    960
369 atgtgggcgg agcttaaagc cctggccgcc tgcagggacg gccgggtgca ccgggcagca   1020
371 gacccccttgg cggggctaaa ggacctcaag gaggtccggg gcctcctcgc caaggacctc   1080
373 gccgtcttgg cctcgaggga ggggctagac ctcgtgcccg gggacgaccc catgctcctc   1140
375 gcctacctcc tggacccctc caacaccacc cccgagggggg tggcgcggcg ctacggggggg   1200
377 gagtggacgg aggacgccgc ccaccgggcc ctcctctcgg agaggctcca tcggaacctc   1260
379 cttaagcgcc tcgaggggga ggagaagctc ctttggctct accacgaggt ggaaaagccc   1320
381 ctctcccggg tcctggccca catggaggcc accggggtac ggcgggacgt ggcctacctt   1380
383 caggcccttt ccctggagct tgcggaggag atccgccgcc tcgaggagga ggtcttccgc   1440
385 ttggcgggcc acccccttcaa cctcaactcc cgggaccagc tggaaagggt gctctttgac   1500
387 gagcttaggc ttcccgcctt ggggaagacg caaaagacag gcaagcgctc caccagcgcc   1560
389 gcggtgctgg aggccctacg ggaggcccac cccatcgtgg agaagatcct ccagcaccgg   1620
391 gagctcacca agctcaagaa cacctacgtg gaccccctcc caagcctcgt ccacccgagg   1680
393 acgggccgcc tccacacccg cttcaaccag acggccacgg ccacggggag gcttagtagc   1740
395 tccgaccccca acctgcagaa catccccgtc cgcacccect tgggccagag gatccgccgg   1800
397 gccttcgtgg ccgaggcggg ttgggcgttg gtggccctgg actatagcca gatagagctc   1860
399 cgcgtcctcg cccacctctc cggggacgaa aacctgatca gggtcttcca ggaggggaag   1920
401 gacatccaca cccagaccgc aagctggatg ttcggcgtcc ccccggaggc cgtggacccc   1980
403 ctgatgcgcc gggcggccaa gacggtgaac ttcggcgtcc tctacggcat gtccgcccat   2040
405 aggctctccc aggagcttgc catcccctac gaggaggcgg tggcctttat agagcgctac   2100
407 ttccaaagct tccccaaggt gcgggcctgg atagaaaaga ccctggagga ggggaggaag   2160
409 cggggctacg tggaaaccct cttcggaaga aggcgctacg tgcccgacct caacgcccgg   2220
411 gtgaagagcg tcagggaggc cgcggagcgc atggccttca acatgcccgt ccagggcacc   2280
413 gccgccgacc tcatgaagct cgccatggtg aagctcttcc cccgcctccg ggagatgggg   2340
415 gcccgcatgc tcctccaggt ccacgacgag ctcctcctgg aggcccccca agcgcgggcc   2400
417 gaggaggtgg cggctttggc caaggaggcc atggagaagg cctatcccct cgccgtgccc   2460
419 ctggaggtgg aggtggggat ggggggaggac tggctttccg ccaagggtta g           2511
422 <210> SEQ ID NO: 10
423 <211> LENGTH: 26
424 <212> TYPE: DNA
425 <213> ORGANISM: Artificial Sequence
427 <220> FEATURE:
428 <223> OTHER INFORMATION: Synthetic
430 <400> SEQUENCE: 10
431 caggaggagc tcgttgtgga cctgga                                          26
```

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/777,430A

DATE: 03/25/2002
TIME: 10:43:01

Input Set : **A:\04912-1.txt**
Output Set: **N:\CRF3\03252002\I777430A.raw**

```
L:2453 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:31
L:2477 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:32
L:2513 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:34
L:2543 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:36
L:2567 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:37
```